

TILAK MAHARASHTRA VIDYAPEETH, PUNE
MASTER OF SCIENCE (M.SC) IN COMPUTER APPLICATIONS
EXAMINATION : JUNE / JULY- 2022
SEMESTER - II

Sub: Data Mining & Data Warehousing - (MSC 206-19)

Date : 05/07/2022

Total Marks : 60

Time: 2.00 pm to 04.30 pm

Instruction:

1. All questions are compulsory unless and otherwise stated.
2. Bold figures to the right of every question are the maximum marks for that question.
3. Candidates are advised to attempt questions in order.
4. Answers written illegibly are likely to be marked zero.
5. Use of scientific calculators, Log tables, Mollier Charts is allowed.
6. Draw neat and labelled diagrams wherever necessary.

Q.1 Answer the following in 2-3lines. (Any 5) (10)

1. Define Post pruning
2. Define decision tree.
3. Define CUBES
4. What are frequent data pattern?
5. Why is the need of doing machine learning?
6. What is KDD?
7. Define data mart.

Q.2 Answer the following in short. (Any 4) (20)

1. Describe K-mean algorithm.
2. Explain difference between data integration and data transformation.
3. Describe cluster analysis in detail
4. Describe dimensional modeling? Explain benefits and limitations of dimensional modeling.
5. Explain decision tree. Why it is mainly used?
6. Explain smoothing technique.

Q.3 Answer the following in detail. (Any 3) (30)

1. Describe hierarchical method of clustering.
2. Describe data discretization in detail
3. Explain Wavelet transform in detail.
4. What is data warehousing? What can a data warehouse do? List five common data warehousing applications.
5. Describe the steps followed in data mining process.